A peer-reviewed open-access journal
NeoBiota
Advancing research on alien species and biological invasions

NeoBiota 105: 319–336 (2026)
DOI: 10.3897/neobiota.105.180959

Methods

Image-based recognition using advanced neural networks can aid surveillance of *Agrilus* jewel beetles

Valerio Caruso[1*], Hossein Shirali[2*], Christophe Bouget[3], Pierfilippo Cerretti[4], Gianfranco Curletti[5], Maarten de Groot[6], Eva Groznik[6,7], Jerzy M. Gutowski[8], Christian Pylatiuk[2], Radosław Plewa[9], Alain Roques[10], Aurélien Sallé[11], Jon Sweeney[12], Kate Van Rooyen[12], Lorenz Wührl[2], Davide Rassati[1]

1 *Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020, Legnaro (PD), Italy*
2 *Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), 76344 Eggenstein-Leopoldshafen, Germany*
3 *Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE), EFNO, Nogent-sur-Vernisson, France*
4 *Department of Biology and Biotechnologies "Charles Darwin", Sapienza University of Roma, Roma, Italy*
5 *Museo Civico di Storia Naturale, Parco Cascina Vigna, 10022 Carmagnola, Italy*
6 *Slovenian Forestry Institute, Department of Forest Protection, Večna Pot 2, 1000 Ljubljana, Slovenia*
7 *Biotechnical Faculty, University of Ljubljana, Jamnikarjeva ulica 101, 1000 Ljubljana, Slovenia*
8 *Department of Natural Forests, Forest Research Institute, 6 Park Dyrekcyjny St., 17-230 Białowieża, Poland*
9 *Department of Forest Protection, Forest Research Institute, Sękocin Stary, 3 Braci Leśnej St., 05-090 Raszyn, Poland*
10 *Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE), UR 0633, Zoologie Forestière, 45075, Orléans, France*
11 *Laboratoire Physiologie, Ecologie et Environnement (P2E), Université d'Orléans, UR 1207, USC-INRAE 1328, 1 rue de Chartres, Orléans 45067, France*
12 *Atlantic Forestry Centre, Canadian Forest Service, Natural Resources Canada, 1350 Regent St., Fredericton, NB, E3C 2G6, Canada*

Corresponding authors: Valerio Caruso (valerio.caruso@unipd.it); Hossein Shirali (hossein.shirali@kit.edu)

Abstract

The genus *Agrilus* includes two species, *Agrilus planipennis* and *A. anxius*, that are of particular phytosanitary concern and that are regulated by the European Union legislation. This implies that phytosanitary agencies of all EU countries are obliged to establish specific surveillance programmes to verify the absence of these species from their territory. These activities commonly consist of the use of green-coloured traps, which are, however, attractive not only for *A. planipennis* and *A. anxius*, but also for a wide range of other *Agrilus* species. For this reason, much time and expertise is required to sort and identify specimens to species, impeding an efficient rapid response. In this study, we tested the efficacy of the Entomoscope, a low-cost, open-source photomicroscope that uses high-resolution digital imaging and allows a pre-trained Convolutional Neural Networks (CNN) model to accurately detect, image and classify insect specimens, for automatic identification of 13 *Agrilus* species, including *A. planipennis* and *A. anxius*. We benchmarked models from three different CNN architectures and selected YOLOv8l as the most robust performer; this model achieved a Top-1 accuracy of 90.2% on a "real-world" test set (i.e. a dataset simulating real surveillance conditions). For most species, including *A. planipennis* and *A. anxius*, either no errors or only a few errors were made, whereas for a few native species, misidentifications were more common. These results provided proof of concept for an AI-driven surveillance system that can strongly aid in surveillance activities of *Agrilus* species.

Key words: *Agrilus anxius*, *Agrilus planipennis*, bronze birch borer, deep learning, early-detection, emerald ash borer, Entomoscope

OPEN ACCESS

Academic editor: Victoria Lantschner
Received: 1 December 2025
Accepted: 22 January 2026
Published: 24 February 2026

Citation: Caruso V, Shirali H, Bouget C, Cerretti P, Curletti G, de Groot M, Groznik E, Gutowski JM, Pylatiuk C, Plewa R, Roques A, Salle A, Sweeney J, Van Rooyen K, Wührl L, Rassati D (2026) Image-based recognition using advanced neural networks can aid surveillance of *Agrilus* jewel beetles. NeoBiota 105: 319–336. https://doi.org/10.3897/neobiota.105.180959

* These authors contributed equally to this work.

Introduction

The constant increase of global trade over the last hundred years, combined with deliberate plant introductions in the past and ongoing climate changes, has facilitated the movement amongst continents and the establishment likelihood of an increasing number of insects (Brockerhoff and Liebhold 2017; Pureswaran et al. 2022; Fenn-Moltu et al. 2023; Isitt et al. 2024). The genus *Agrilus* (Coleoptera, Buprestidae) is amongst the insect taxa most favoured by these processes and is commonly transported to and introduced into new regions. This genus includes 3341 species and, therefore, is one of the most species-rich genera of animals in the world (Jendek and Grebennikov 2023). More than 30 species have already been introduced and established outside their native range (Ruzzier et al. 2023), including the emerald ash borer *A. planipennis* Fairmaire, 1888 which has caused massive ecological and economic damage in North America (Kovacs et al. 2010; Klooster et al. 2018), Russia and Ukraine (Orlova-Bienkowskaja et al. 2020). For these reasons, the development of tools and strategies for the early detection of accidentally introduced *Agrilus* species was identified as a research priority to trigger rapid response and reduce potential impacts in the invaded areas.

Amongst the numerous tools developed for *Agrilus* surveillance programmes, baited or unbaited green traps set up at entry points and in their surrounding forests are currently adopted by several phytosanitary agencies worldwide (Evans et al. 2020; Imrei et al. 2020; Silk et al. 2020; Dodds et al. 2024; Duan et al. 2024; Santoiemma et al. 2024a). These traps were primarily developed, based on laboratory and field studies targeting *A. planipennis* (Crook et al. 2009; Francese et al. 2010; Poland et al. 2019), but are also attractive for a wide range of other species within the genus *Agrilus* (Rassati et al. 2019; Cavaletto et al. 2020; Kuhn et al. 2024; Le Souchu et al. 2024; Santoiemma et al. 2024b, 2025). This allows practitioners to survey multiple species simultaneously, but much time and expertise are required to sort and identify specimens to species due to both the low interspecific morphological differentiation (Kelnarova et al. 2019) and the very high species richness of the genus (Jendek and Grebennikov 2023), impeding an efficient, rapid response (Lyal and Miller 2020). Thus, novel technologies that can help to improve the efficiency of species identification are promptly needed.

Artificial intelligence (AI) systems are being increasingly adopted in entomology for various applications (Teixeira et al. 2023; Hartbauer 2024). Amongst them, AI is used to enhance the accuracy and speed of insect identification in the laboratory and in the field (De Cesaro Júnior and Rieder 2020; Gao et al. 2024; Hartbauer 2024). Convolutional neural networks (CNNs) trained on image datasets have been shown to reliably classify insects at the family, genus or even species level (Valan et al. 2019; Ärje et al. 2020; Hansen et al. 2020; Wührl et al. 2022; Tannous et al. 2023; Lertrusdachakul et al. 2025; Marais et al. 2025). Exploiting this technology in surveillance activities for *Agrilus* species would, however, require that laboratories and plant health inspectors are equipped with an affordable device that can capture images of trapped specimens and, subsequently, automatically identify them to species. The Entomoscope, recently developed by the Karlsruhe Institute of Technology (KIT) (Wührl et al. 2024), might satisfy these needs. It is a low-cost, open-source photomicroscope that uses high-resolution digital imaging and allows a pre-trained CNN to accurately detect, image and classify insect specimens (Wührl et al. 2024). The Entomoscope has been tested so far on parasitoid wasps (Shirali et al. 2024), but its potential application in *Agrilus* surveillance has not yet been investigated.

In this study, we evaluated the Entomoscope's efficacy for *Agrilus* identification, with a specific focus on model robustness for "real-world" deployment, i.e. performance on data that mimic actual surveillance workflows, with new trap collections acquired under realistic variation in lighting, specimen condition and operator handling. First, we benchmarked multiple state-of-the-art CNN architectures to select the top performers. Second, we went beyond standard "laboratory" validation, i.e. accuracy measured on randomly shuffled data acquired under controlled and uniform conditions and tested these models against a dataset designed to simulate a surveillance process, based on traps. We compared this "real-world" performance to the often-optimistic results from standard validation on randomly-shuffled data. Finally, we tested the model's ability to reject "unknown" species not included in the training data, with the objective of developing a truly robust, deployable AI tool to be used along with traps for *Agrilus* surveillance.

Materials and methods

Agrilus specimens

The dataset for this study comprised specimens from 13 *Agrilus* species, including 11 native species to Europe and two non-native species of high phytosanitary concern: *A. planipennis* (native to Asia) and *A. anxius* (native to North America) (Suppl. material 1: table S1, Fig. 1). The native species were selected as they are commonly collected in trapping studies carried out in Europe targeting *Agrilus* spp. (Le Souchu et al. 2024; Santoiemma et al. 2024b, 2025), whereas the non-natives were selected because they are regulated by the European Union legislation and surveillance is mandatory in EU countries under Regulation (EU) 2019/2072 (EFSA et al. 2020a, b). Specimens were collected in Canada, France, Poland and Slovenia (Suppl. material 1: table S1), using green multi-funnel traps. The preservation methods in the trap-collecting cups varied by location and included: a 2:1 mixture of propylene glycol and water (Slovenia); a 1:1 mixture of ethylene glycol and water (Poland); a 2:1 mixture of monopropylene glycol and water with a drop of liquid dish detergent or kept dry, but integrated with a section of mesh impregnated with α-cypermethrin insecticide (Storanet®, BASF Pflanzenschutz Deutschland, Germany) (France); and a saturated salt water solution plus a drop of liquid dish detergent to reduce surface tension (Canada). Specimens were preserved in ethanol until identification. Species-level taxonomic identifications were initially performed by Eva Groznik, Maarten de Groot, Jerzy M. Gutowski, Alain Roques, Aurélien Sallé and Kate Van Rooyen, based on morphological traits, keys and other reference materials (Schaefer 1950; Farrugia 2007; Paiero et al. 2012) and then confirmed by Gianfranco Curletti.

Image acquisition

A standardised imaging protocol was established using the "plug-in" version of the Entomoscope (Wührl et al. 2024). Two Entomoscope (Suppl. material 1: fig. S1) units were utilized: one at the University of Padova, DAFNAE department and another at the Entomology Museum of La Sapienza University

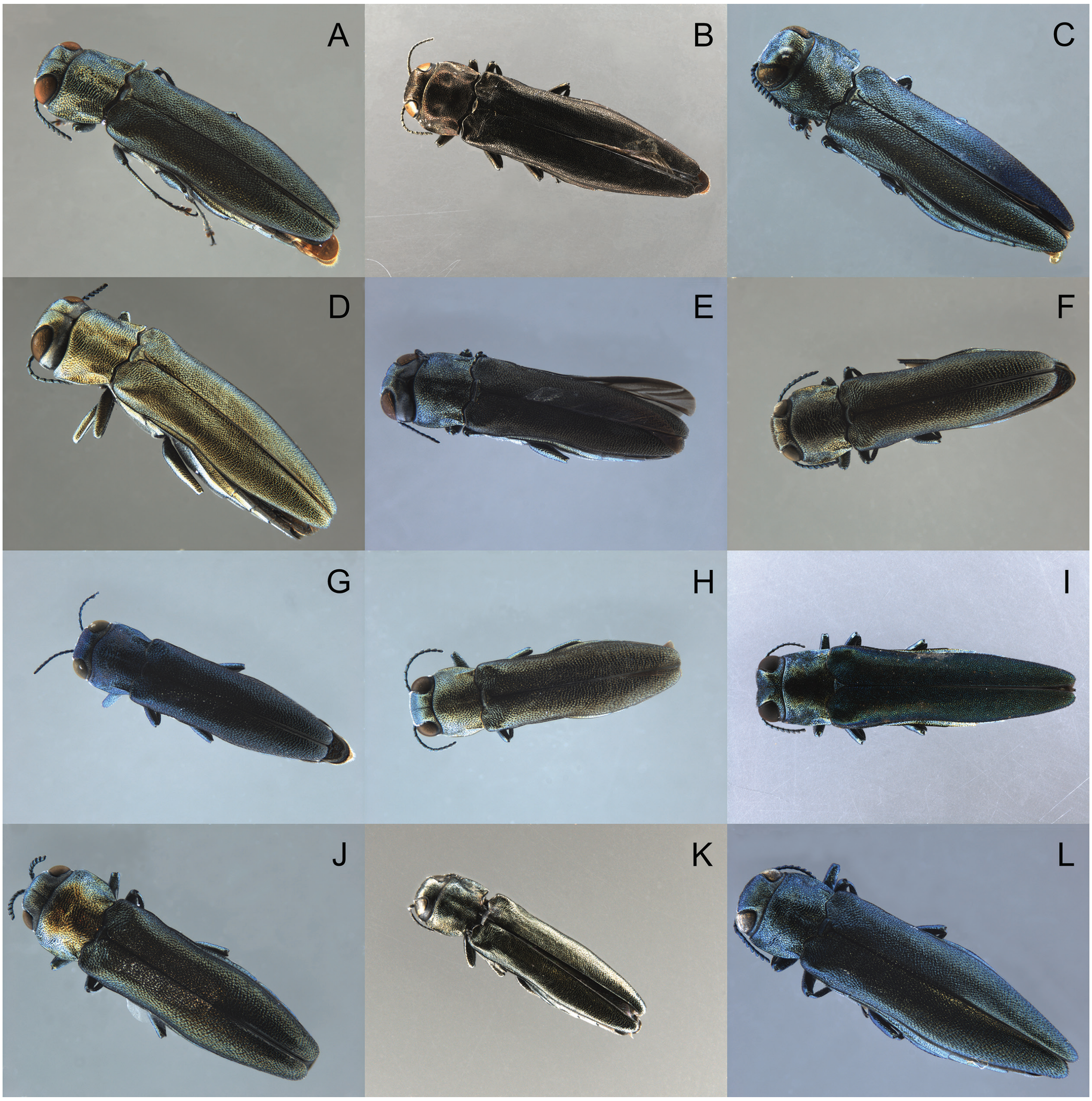

Figure 1. Examples of pictures taken using the Entomoscope to some of the *Agrilus* specimens used during the training process. **A.** *A. angustulus*; **B.** *A. anxius*; **C.** *A. cuprescens*; **D.** *A. graminis*; **E.** *A. hastulifer*; **F.** *A. laticornis*; **G.** *A. obscuricollis*; **H.** *A. olivicolor*; **I.** *A. planipennis*; **J.** *A. pratensis*; **K.** *A. sulcicollis*; **L.** *A. viridis*. Note: the pictures are not to scale relative to each other.

of Rome. To simulate realistic conditions faced by phytosanitary inspectors, specimens were not manipulated or positioned (or at least to a minimal extent for the identification process). Instead, they were photographed in their natural preservation state (i.e. "death position") and without distinguishing between sexes. Each specimen was photographed from two to five standard points of view to ensure comprehensive morphological coverage: dorsal, laterodorsal, lateral, lateroventral and ventral (Suppl. material 1: fig. S2). For image capture, specimens were submerged in 70% ethanol to minimise distortion effects. All images were acquired using the Entomoscope fixed, standard-intensity lighting. To create a single, fully-focused image for each angle, a focus stacking technique using Helicon Focus (v.8.2.2) was employed.

Dataset preparation and evaluation strategy

Our evaluation strategy was designed to assess model robustness against the challenges of surveillance. The dataset consisted of 14 classes: 13 *Agrilus* species and one "Background" class. The "Background" class consisted of 89 images of the empty Entomoscope setup (empty Petri dish, with and without ethanol) taken under various lighting conditions to simulate non-target images. Across the 13 species included in model development, the number of specimens per species ranged from 11 to 59 (mean ± SD = 36.1 ± 13.8; see Suppl. material 1: table S1), for a total of 469 specimens. Each specimen contributed between two and six images (five points of view), resulting in multiple images per individual for model training and evaluation. The entire dataset was first partitioned at the specimen level (to prevent data leakage). This partitioning was stratified by class (to maintain the distribution of all 14 classes), but ordered by acquisition time (to create a realistic temporal split). Two sets of images were created: i) the development set, including the first 85% of images from the ordered, stratified dataset. This set represents the data available for all model training and validation; ii) the held-out "real-world" set, including the final 15% of images from the ordered, stratified dataset. As these images were collected later, they may contain subtle variations in lighting, operator setup or new specimen batches. This set served as our challenging OOD test set, simulating the data drift common in surveillance (often termed an Out-Of-Distribution or OOD dataset).

Our evaluation followed a two-stage process. Stage 1 consisted of the model benchmarking. In particular, we trained 11 model variants (from You Only Look Once [YOLOv8, YOLO11] and EfficientNet families) using an ordered 70/15 (a total of 85% of the whole set) split of the Development Set. This means that the first 70% of the Development Set was used for training and the 15% was used as a validation set to select the champion models. The models from this stage were also evaluated on the OOD Set to establish a baseline performance from a simple, ordered training strategy. Stage 2 consisted in the robustness evaluation. The top-performing models were advanced for further evaluation. We performed a randomly-shuffled, stratified 5-fold cross-validation (CV) only on the 85% Development Set. The folds were partitioned at the specimen ID level to prevent data leakage. For each of the five trained folds, we performed two distinct evaluations: i) Standard Validation (IID): the performance was measured on the held-out validation fold from the cross-validation. This represents a standard test on randomly-shuffled data (Independent and Identically Distributed or IID); ii) Real-World Validation (OOD). The exact same model was then tested against the 15% (remaining 15% not used in training and validation process) OOD Set to measure robustness to data drift. The final reported metrics for our recommended methodology were the average and standard deviation across these five folds. This process validated the training protocol. To test the model's rejection capabilities, one additional class was created and used only for testing the "unknown" species class (see below). This class included two species not in the training set, *A. biguttatus* and *A. convexicollis*.

Model architecture

We evaluated and compared three distinct state-of-the-art Convolutional Neural Network (CNN) architectures for our fine-grained classification task, i.e. YOLOv8, YOLO11 and EfficientNet. All models were initialised with weights, pre-trained on the ImageNet dataset (Deng et al. 2009) and subsequently fine-tuned on our *Agrilus* dataset. The Ultralytics YOLO (Limberg et al. 2022; Jocher et al. 2023; Khanam and Hussain 2024) framework provides a suite of models optimised for various computer vision tasks. Each model version is available in five sizes (nano, small, medium, large and extra-large), which offer a trade-off between speed and accuracy. For this study, we selected their dedicated classification models. These models utilise the powerful and highly efficient YOLO backbone for feature extraction, but employ a standard classification head, making them distinct from the object detection variants and ideally suited for image classification. Their inclusion allows us to evaluate architectures that provide a state-of-the-art balance between inference speed and accuracy, a critical consideration for future deployment in high-throughput surveillance systems. To benchmark the YOLO models against a well-established, pure classification architecture, we included EfficientNet (Tan and Le 2019). This model family is renowned for achieving high accuracy through a principled compound scaling method and serves as a strong baseline for classification performance. We employed the EfficientNetV2 large variant (Tan and Le 2021). This comparative approach allows for a robust evaluation, pitting models optimised for speed and efficiency (YOLO series) against an architecture designed to maximise classification accuracy (EfficientNet).

Model training and implementation

All models were trained on the HAICORE high-performance computing cluster at the Karlsruhe Institute of Technology (KIT), utilising nodes equipped with NVIDIA A100-SXM4-40GB GPUs. The software environment consisted of Python (v. 3.10) and PyTorch (v. 2.0.1) with the Ultralytics framework for the YOLO models and TensorFlow (v. 2.10) with Keras for the EfficientNetV2 models. We employed transfer learning for all models, initialising them with weights pre-trained on the ImageNet dataset (Deng et al. 2009) to accelerate training and improve generalisation. Input images were resized to the input dimensions required by each architecture (e.g. 640×640 for YOLOv8, 480×480 for EfficientNetV2L) and standard data augmentation techniques (e.g. random flips, rotations, colour adjustments, erasing) were applied. To ensure robust generalisation and prevent overfitting, we employed a consistent optimisation strategy. All models were fine-tuned using the AdamW optimiser (Loshchilov and Hutter 2017) and included dropout regularisation (rate = 0.3). An early stopping mechanism was used, halting training if the validation loss did not improve for 15 consecutive epochs, with a maximum limit of 150 epochs. To address the inherent class imbalance in the dataset, we applied class weights to the loss function. For the EfficientNetV2 models, a categorical focal cross-entropy loss was used to further prioritise difficult-to-classify examples, while the YOLO models used their framework's standard weighted loss function. The training and validation curves, as well as accuracy curves are shown in Fig. 2. The learning dynamics showed similar patterns across folds, with consistent loss decay and accuracy stabilisation; therefore, one representative example (fold 4) is shown here. All the images used for training and testing, as well as the scripts employed in the classification pipeline, are available in the Zenodo (www.zenodo.org) repository.

Performance evaluation metrics

Model performance was assessed using a comprehensive set of metrics, evaluated at two distinct stages. In stage 1 (benchmarking), models were ranked, based on

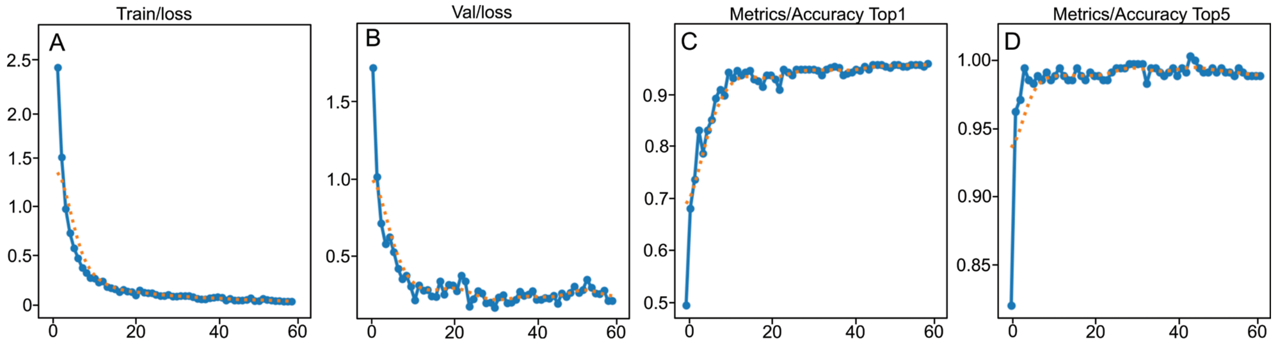

Figure 2. Representative training and validation curves for the YOLOv8l model during one of the cross-validation folds. The plots show training loss (**A**) and validation loss (**B**) decreasing over epochs, indicating effective learning; Top-1 accuracy (**C**) and Top-5 accuracy (**D**) on the validation set increasing and stabilising, confirming robust convergence without overfitting. Blue full line = results; orange dashed line = smooth.

Top-1 validation accuracy. Their performance on the OOD Set was also recorded. In stage 2 (robustness), the 5-fold CV models were evaluated on both the IID validation fold and the OOD test set, using Top-1 Accuracy and Weighted F1-Score. The final performance of the methodology was reported as the average of the 5-fold CV results on the OOD set. A full classification report and an aggregated confusion matrix, summing the predictions from all five folds, were generated. To test the model's ability to handle species not included in the training process, we added an additional class named "unknown". In fact, our model does not include an "out-of-class" mechanism to exclude images that do not belong to any trained class, so it still tries to associate the image with the most similar species. When the image is very different from those used for training, the confidence level (max_conf) of the prediction is typically low. In our case, we set this confidence level to 0.5, meaning that the model will ignore all predictions with less than 50% confidence. This threshold was adopted following the standard convention, as it represents the neutral decision boundary in probability-based classification tasks. To test this mechanism, we used 19 images belonging to two native species that were not used for the training procedure (i.e. 8 images belonging to 2 specimens of *A. biguttatus* and 11 images belonging to 3 specimens of *A. convexicollis*).

Model explainability

To gain insight into the models' decision-making process, we employed the Eigen-CAM explainability technique (Muhammad and Yeasin 2020). Eigen-CAM produces saliency maps that highlight the image regions most influential for a given classification, allowing for a qualitative assessment of whether the models focus on relevant morphological features.

Results

Stage 1: model benchmarking

The 11 model variants were trained on an ordered 70% training set and ranked by their Top-1 accuracy on the ordered 15% validation set (Table 1). The EfficientNetV2L and YOLOv8l emerged as the top performers, tying with an identical validation accuracy of 88.9% (0.8885). These two models were selected for the comprehensive Stage 2 evaluation. When tested on the OOD Set, the models

Table 1. Benchmarking of 11 model variants (YOLOv8l and EfficientNetV2L). Reports Validation Top-1 (ordered split) and, on the held-out real-world OOD test set, Test Top-1 and Test Weighted F1. Values in are reported as point estimates with 95% confidence intervals (Wilson) for Top1 accuracy, to quantify uncertainty due to finite validation/test sample sizes.

	Validation Top1 accuracy (95% CI)	Test Top1 accuracy (95% CI)	Test Weighted F1 score
Yolov8n	0.8500 (0.8047–0.8850)	0.8100 (0.7618–0.8504)	0.7900
Yolov8s	0.8852 (0.8446–0.9163)	0.8100 (0.7618–0.8504)	0.7800
Yolov8m	0.8787 (0.8373–0.9107)	0.8200 (0.7726–0.8593)	0.8000
Yolov8l	0.8885 (0.8483–0.9191)	0.8500 (0.8052–0.8860)	0.8300
Yolov8x	0.8852 (0.8446–0.9163)	0.8300 (0.7834–0.8683)	0.8100
Yolo11n	0.8623 (0.8191–0.8965)	0.8100 (0.7618–0.8504)	0.7800
Yolo11s	0.8525 (0.8083–0.8879)	0.7800 (0.7297–0.8232)	0.7600
Yolo11m	0.8787 (0.8373–0.9107)	0.8000 (0.7511–0.8413)	0.7800
Yolo11l	0.8400 (0.7940–0.8763)	0.7600 (0.7086–0.8048)	0.7400
Yolo11x	0.8426 (0.7975–0.8792)	0.7700 (0.7191–0.8140)	0.7500
EfficientNetV2L	0.8885 (0.8483–0.9191)	0.8500 (0.8052–0.8860)	0.8200

from this simple, ordered training strategy performed modestly. The top-performing YOLOv8l achieved a Top-1 accuracy of 85.0% (Table 1), establishing a baseline for our subsequent robustness tests.

Stage 2: Robustness evaluation of standard vs. real-world testing

The top two models, YOLOv8l and EfficientNetV2L, underwent a 5-fold CV. Both models achieved near-perfect accuracy on the IID validation folds, with YOLOv8l averaging 97.9 ± 1.4% and EfficientNetV2L averaging 97.2 ± 2.4% (Table 2, Fig. 3). However, when these same models were tested against our 15% OOD Set, two key findings emerged. First, a performance gap was observed. The YOLOv8l model's accuracy, for example, was 97.9% on the IID fold, but 90.2% on the OOD set, a difference of 7.7 percentage points. Second, the models trained with the robust, randomly-shuffled CV methodology (Stage 2) achieved higher accuracy on the OOD set than the models from the simple, ordered training (Stage 1). The CV-trained YOLOv8l achieved an average Top-1 accuracy of 90.2 ± 1.9%,

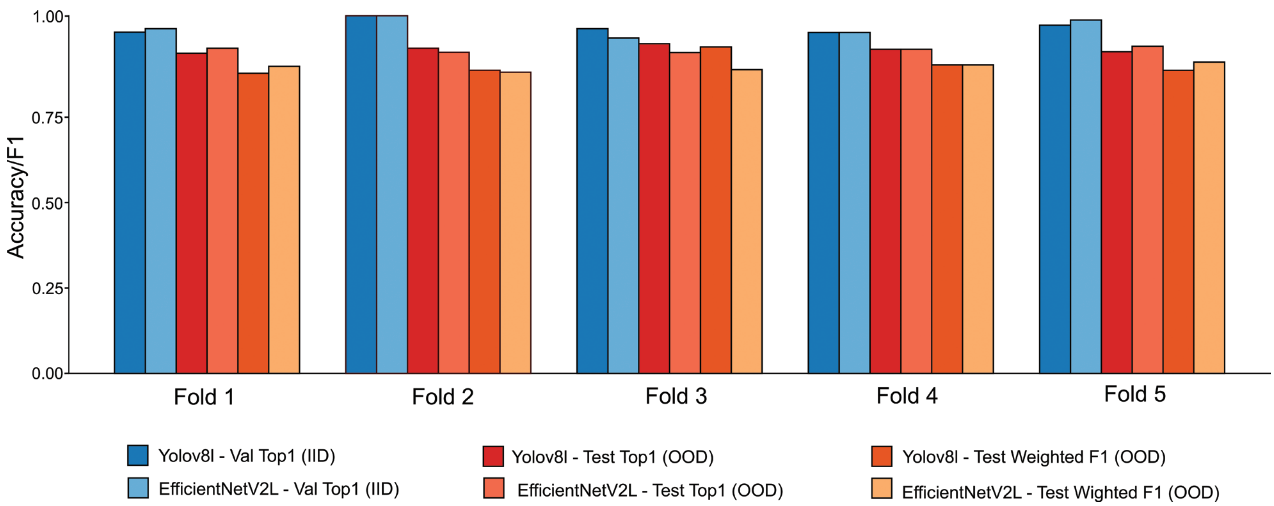

Figure 3. Comparison of Top-1 Accuracy between Standard Validation (IID) and "Real-World" Validation (OOD) for YOLOv8l and EfficientNetV2L models. Bars show accuracy across five folds, illustrating the performance gap between IID and OOD evaluations.

Table 2. Results (%) of the 5-fold cross-validation for YOLOv8l and EfficientNetV2L models.

Model	Metric	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean ± SD
EfficientNetV2L	Validation Top1 Accuracy	97.0	100.0	94.0	96.0	99.0	97.2 ± 2.4
EfficientNetV2L	Test Top1 Accuracy	90.0	88.0	88.0	91.0	92.0	89.8 ± 1.7
EfficientNetV2L	Test Weighted F1	88.0	86.0	87.0	89.0	90.0	88.0 ± 1.6
YOLOv8l	Validation Top1 Accuracy	96.6	100.0	98.0	96.6	98.5	97.9 ± 1.4
YOLOv8l	Test Top1 Accuracy	88.0	90.0	93.0	91.0	89.0	90.2 ± 1.9
YOLOv8l	Test Weighted F1	86.0	87.0	92.0	89.0	87.0	88.2 ± 2.4

compared to the 85.0% from the Stage 1 ordered-split model. On this OOD test, the YOLOv8l methodology had a higher average Top-1 accuracy, with 90.2 ± 1.9% and a weighted F1-score of 88.2 ± 2.4%. This was consistently higher than the EfficientNetV2L (89.8 ± 1.7% Top-1, 88.0 ± 1.6% F1).

Final YOLOv8l model performance (5-Fold Average)

Based on its higher robustness and performance, YOLOv8l was selected as our final recommended methodology. The trained CNN correctly assigned the highest percentage of correspondence to the correct species 100% of times for seven species, i.e. the non-native *A. anxius* and *A. planipennis* and the native *A. angustulus*, *A. betuleti*, *A. cuprescens*, *A. graminis* and *A. olivicolor* (Fig. 4). For four of the other species, the trained CNN assigned the highest percentage of correspondence to the correct species between 86% and 98% of times (Fig. 4). In particular, the native *A. laticornis* was identified with a 92% accuracy (respectively 4% and 5% of the pictures with *A. laticornis* were misidentified with *A. angustulus* and *A. graminis*), the native *A. obscuricollis* with a 95% accuracy (3.5% misidentified as *A. olivicolor* and 1.5% as *A. hastulifer*), the native *A. sulcicollis* with 86.0% accuracy (9% misidentified as *A. laticornis*, 1% as *A. planipennis* and 3% as *A. angustulus*) and the native *A. viridis* with a 98% accuracy (2% misidentified as *A. graminis*). The *Agrilus* species for which a misidentification was common were *A. hastulifer* and *A. pratensis*. For *A. hastulifer*, the trained CNN assigned the highest percentage of correspondence to *A. graminis* 65% of the times and to *A. viridis* 21% of the times (Fig. 4), making *A. hastulifer* the class with the lowest percentage of correct classification (9%). For *A. pratensis*, the trained CNN assigned the highest percentage of correspondence to the wrong species 58% of the times to *A. obscuricollis* and 2% times to *A. hastulifer* (Fig. 4). Full per-class metrics (Precision, Recall, F1, Support; mean ± SD over five folds) are provided in Suppl. material 1: table S2.

Identification of out-of-distribution species

The model's rejection capability was tested on the OOD set. Averaged across the five folds, the YOLOv8l methodology correctly classified 100% of the 19 images from the "unknown" class (*A. biguttatus* and *A. convexicollis*) as unknown species using a 0.5 confidence threshold.

Model explainability

Eigen-CAM analysis of the trained YOLOv8l models revealed that the classification decisions relied on specific morphological regions. For example, the model consistently focused on the distinctive shape and texture of the elytra in most specimens in

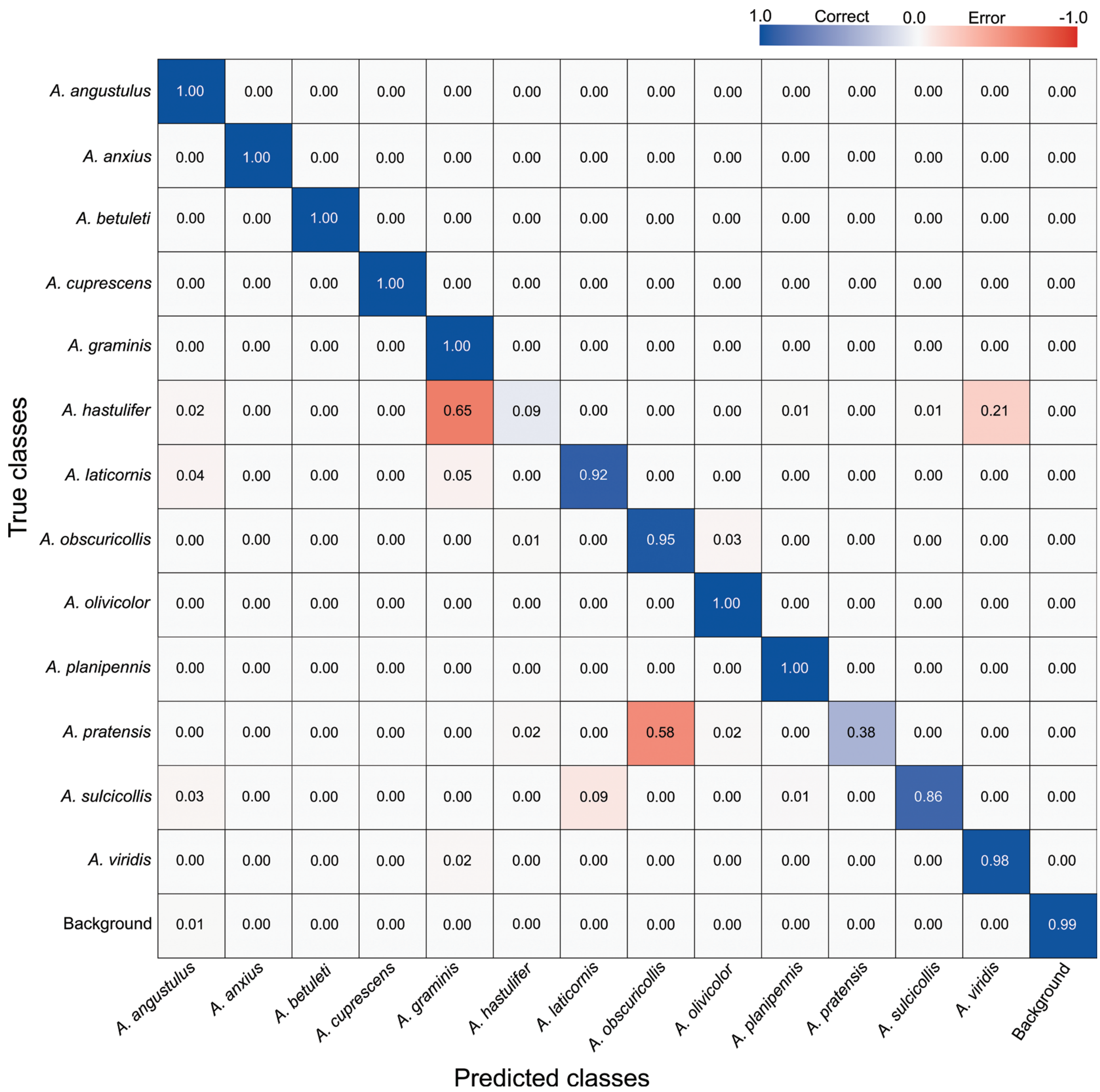

Figure 4. Aggregated confusion matrix for YOLOv8l on the held-out real-world (OOD) test set. The matrix represents the sum of all predictions from the five cross-validation folds across 14 classes, showing classification accuracy per class and highlighting the main sources of misclassification. True classes (Y axis) indicate the true species identity. Predicted classes (X axis) indicate the species to which the model assigned analysed specimens. The degree of accuracy is reported both in the form of numerical value, showing the percentage of specimens that were assigned to a given species and colour-coded with blue for correct predictions and red for incorrect ones (light blue and red are associated with low degree of prediction, both correct and incorrect).

dorsal and laterodorsal views (*A. angustulus*, Fig. 5A; *A. olivicolor*, Fig. 5C), as well as on head morphology — particularly the pronotal margins, eyes and clypeus (*A. olivicolor*, Fig. 5C). In lateral and ventral images, key features, such as the shape of the abdominal sternites (*A. angustulus*, Fig. 5B; *A. planipennis*, Fig. 5D), the structure of the eyes (*A. planipennis*, Fig. 5D) and, in many cases, the elytra, were especially influential.

Discussion

The genus *Agrilus* includes two species, *A. planipennis* and *A. anxius*, that are of particular phytosanitary concern due to the economic and ecological impacts they can have in the invaded areas (Baranchikov et al. 2008, 2014; Kovacs et al. 2010;

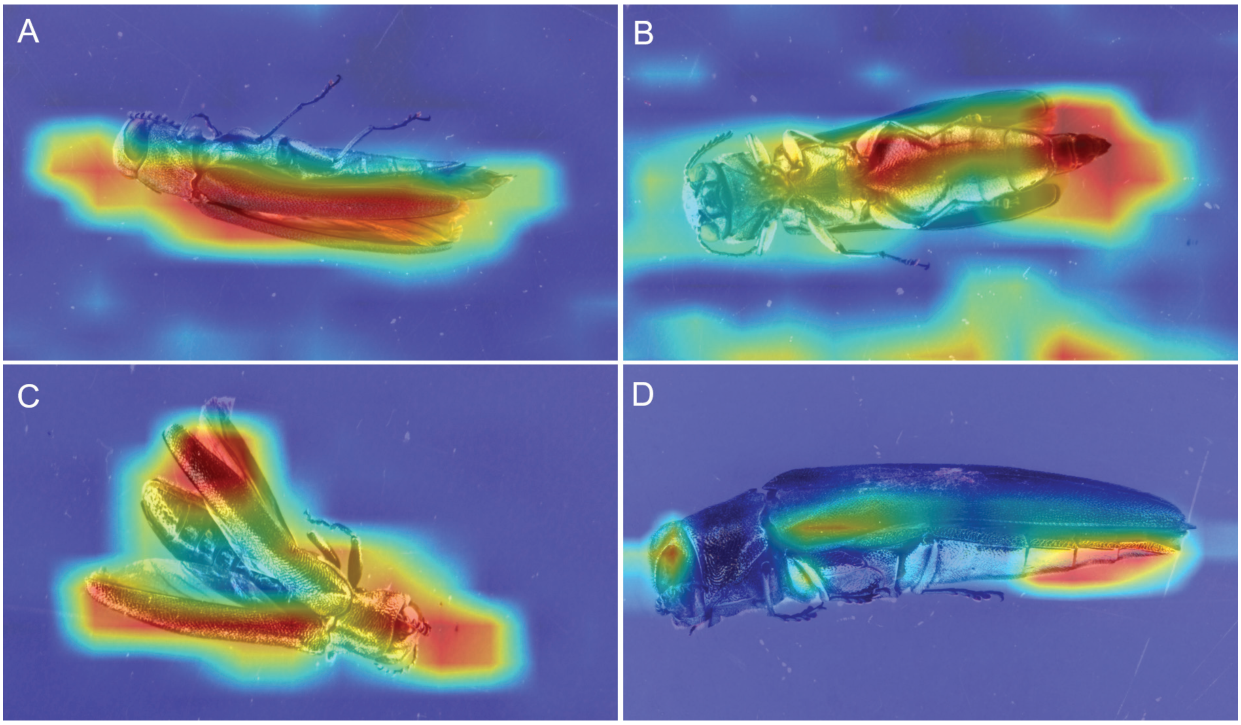

Figure 5. Example of class activation heatmaps obtained for pictures of: **A.** *Agrilus angustulus* (laterodorsal view); **B.** *A. angustulus* (ventral view); **C.** *A. olivicolor* (dorsal view); **D.** *A. planipennis* (lateral view). The heatmaps indicate the image regions most influential for the model's prediction (from red = high influence to blue = low influence).

Klooster et al. 2018; Evans et al. 2020) and, as such, they are regulated by the European Union legislation (EFSA et al. 2020a, 2020b). This implies that phytosanitary agencies of all EU countries are obliged to establish specific surveillance programmes to verify the absence of these species from their territory (Evans et al. 2020). These activities commonly consist of the use of green-coloured traps, which are, however, attractive not only for *A. planipennis* and *A. anxius,* but also for a wide range of other *Agrilus* species (Rassati et al. 2019; Cavaletto et al. 2020; Le Souchu et al. 2024; Santoiemma et al. 2024a, 2024b, 2025). Here, we showed that the Entomoscope, in combination with our validated deep-learning methodology, can strongly aid in discriminating *A. planipennis* and *A. anxius* from other commonly collected native *Agrilus* species, a process that otherwise demands much time and taxonomic expertise.

Our validated YOLOv8l methodology, which averaged 90.2% Top-1 accuracy on the real-world set, serves as a powerful proof-of-concept for a practical triage tool. Its perfect identification (100% F1-score) of the key pests *A. planipennis* and *A. anxius*, as well as its almost perfect identification of the "Background" class, demonstrates its immediate value for phytosanitary agencies. This system could be used to autonomously sort through thousands of specimens, filtering out empty images and flagging only those identified as high-priority pests for expert review. However, the model is not an infallible taxonomist. The aggregated confusion matrix shows that misclassifications still occur amongst a few native species, namely *A. hastulifer* and *A. pratensis*. This is quite surprising, as these species are easily separated using classic taxonomic approaches (Schaefer 1950; Farrugia 2007). This might be due to the characters used by the models. For species represented by a limited number of test specimens, such as *A. cuprescens* and *A. betuleti*, performance metrics should be interpreted with appropriate caution. Nevertheless, the consistently high classification performance metrics observed for these classes indicate that the model is able to reliably capture diagnostic visual features even when trained and evaluated on relatively small sample sizes. As for many other CNNs, the models are only partially based on the same morphological characteristics (such as genitalia, setae and hairs distribution, size etc.) (e.g. Volkovitsh et al. 2020) used by taxonomists, but instead on the arrangement and intensity of pixel clusters which are identified as discriminating features by the algorithm (Goodfellow et al. 2016; Redmon et al. 2016). Our Eigen-CAM analysis supports this, showing that, depending on the position of the specimen, the general shape of the body and the margins of the head, elytra and pronotum or the shape of the abdominal sternites, the forehead and the clypeus were used as main characters to discriminate amongst the tested *Agrilus* species. This can result in accurate classification of species that are difficult to distinguish (i.e. *A. obscuricollis*), but also failures when attempting to discriminate between species that are easily separated using classic taxonomic approaches (e.g. *A. hastulifer* and *A. pratensis*). Overall, these results indicate that the Entomoscope and the current methodology are highly useful for surveillance of *A. planipennis* and *A. anxius*, but are not yet ready to identify all native species collected in traps to a species level.

Our results also demonstrated the ability of the trained CNN to reject non-target images. Our methodology correctly assigned 100% of correspondence to the class "unknown" when identifying specimens belonging to species not used in the training process (*A. biguttatus* and *A. convexicollis*). Several trapping studies showed that many *Agrilus* species can be collected when using green-coloured traps (e.g. Santoiemma et al. 2024a, 2024b). These species might include native species present at low density or non-native species not initially targeted. In this scenario, the capacity of the trained methodology to discriminate between known and unknown species can alert the phytosanitary personnel of the presence of a potential non-native species that should be further scrutinised via classic morphological or molecular approaches.

Our results also indicate that the method of validation is critical for determining real-world-viability. Our study demonstrates a significant "validation gap" between IID tests and our more realistic, OOD dataset. While our top models achieved near-perfect accuracy (97–98%) in the "laboratory" setting, this performance dropped by over seven percentage points when faced with data simulating new collection events. This finding strongly suggests that studies reporting accuracy from simple, random splits may be dangerously optimistic and that testing against OOD data is essential for validating deployable biosecurity tools. Furthermore, our study provides a key insight into how to train a robust model. The model from our initial benchmark, trained on a simple ordered split, achieved a respectable 85.0% on the OOD set. However, the models trained using a randomly-shuffled, 5-fold CV (on the exact same development data) were demonstrably superior, achieving an average of 90.2%. This proves that a robust training strategy, which exposes the model to more diverse and randomised data combinations, is just as important as the model architecture itself for building a tool that can generalise to new data. This performance gap was not uniform. Interestingly, while the EfficientNetV2L tied for the highest score in our initial benchmark, the YOLOv8l model proved to be more robust, consistently performing better on the challenging OOD set. This suggests that the architectural properties of YOLOv8l may offer better generalisation to the subtle data drift (e.g. in lighting, pose or operator handling) inherent in real-world surveillance.

In addition to domain shift, another important challenge in applied image-based classification is the "open-world" or semantic shift problem. In operational surveillance contexts, traps may collect a large number of non-target species that are not represented in the training dataset. In such cases, closed-set classifiers may erroneously assign these specimens to one of the known classes rather than rejecting them as unknown. We acknowledge that this limitation is intrinsic to supervised classification approaches trained on a finite and predefined set of taxa. Consequently, automated identification should be interpreted as a decision-support and triage tool, rather than as a fully autonomous replacement for expert taxonomic assessment, at least under current technological and data-availability constraints. In practice, this implies that a small subset of specimens, particularly those assigned with lower confidence or representing each predicted species, could be routinely verified by a taxonomic expert. Importantly, the species selected for this study correspond to some of the most frequently collected *Agrilus* species in European trapping programmes, which reduces the practical impact of semantic shift under typical surveillance conditions.

Conclusions

This study provides a robust proof-of-concept for an AI-driven system, combining the Entomoscope with a validated deep-learning methodology, for the surveillance of *Agrilus* jewel beetles. We have demonstrated that our final YOLOv8l methodology is a reliable triage tool, capable of perfectly identifying the high-priority pests *A. planipennis* and *A. anxius*, correctly filtering background noise and effectively rejecting unknown species. In addition, we have presented a validation framework that moves beyond standard "laboratory" accuracy to test for robustness against real-world data drift. This methodology of evaluating on a OOD dataset is critical for bridging the gap between a promising model and a trustworthy, deployable tool for phytosanitary surveillance. From a longer-term perspective, the progressive expansion and diversification of reference image datasets represents the most effective strategy to mitigate the open-world problem. Increasing taxonomic and phenotypic coverage will reduce classification errors and progressively confine rare or unexpected taxa to a marginal "unknown" category with limited operational impact. In parallel, the open-source nature of this technology, combined with the relatively low cost of setting up and operating an Entomoscope, makes AI-based identification systems highly suitable for technology transfer initiatives. These systems could be integrated into training projects for phytosanitary personnel, foresters and environmental agencies, increasing knowledge on AI-based identification technologies and encouraging their active participation in large-scale monitoring efforts. Future work will focus on expanding the species library and integrating this model into a field-ready, rapid-alert system without the need for extensive laboratory infrastructure (Brydegaard et al. 2024; Chiavassa et al. 2024), as well as hybrid approaches that integrate CNN-based image classification with other diagnostic methods, such as molecular tools. A cloud-based or mobile application linked to the Entomoscope could allow users to upload images and receive automated identifications in real time, greatly increasing the accessibility and practical applicability of the system. Such an innovation would not only facilitate rapid species identification in phytosanitary surveillance, but also contribute to broader biodiversity assessment initiatives.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Use of AI

No use of AI was reported.

Funding

This work was supported by the project "Improving Forest Insect Biosecurity in an era of globalization – FORBIO", funded by the European Union-Next Generation EU, Mission 4, Component 1, CUP:C53D23009620001 and B53D23032040001 and partially by the European Union's Horizon Europe Research and Innovation programme under grant agreement No. 101134200 "FORSAID: Forest surveillance with artificial intelligence and digital technologies". MdG and EG were funded by the Administration of the Republic of Slovenia for Food Safety, Veterinary Sector and Plant Protection (Contract No. C2337-23-000026) and Slovenian Research Agency (Research Programme P4-0107); HS and CP were funded by grant #ZF4717901SK9 of the programme Natural, Artificial and Cognitive Information Processing (NACIP) of the Helmholtz-Association, German.

Author contributions

Valerio Caruso: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualisation, Visualisation; Hossein Shirali: Writing – original draft, Methodology, Investigation, Formal analysis, Visualisation; Christophe Bouget: Writing – review and editing; Pierfilippo Cerretti: Writing – original draft, Conceptualisation, Supervision, Funding acquisition; Gianfranco Curletti: Writing – review and editing, Data curation; Maarten de Groot: Writing – review and editing, Resources; Eva Groznik: Writing – review and editing, Resources; Jerzy M. Gutowski: Writing – review and editing, Resources; Pyliatuk: Writing – review and editing, Resources, Methodology; Radosław Plewa: Writing – review and editing, Resources; Alain Roques: Writing – review and editing, Resources; Aurelién Sallé: Writing – review and editing, Resources; Jon Sweeney: Writing – review and editing, Resources; Kate Van Rooyen: Writing – review and editing, Resources; Lorenz Wührl: Writing – review and editing, Methodology; Davide Rassati: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualisation, Supervision, Visualisation, Funding acquisition, Project administration.

Author ORCIDs

Valerio Caruso https://orcid.org/0009-0006-9421-7447
Hossein Shirali https://orcid.org/0009-0005-6884-4263
Christophe Bouget https://orcid.org/0000-0002-5206-7560
Pierfilippo Cerretti https://orcid.org/0000-0002-9204-3352
Maarten de Groot https://orcid.org/0000-0002-5721-6676
Eva Groznik https://orcid.org/0009-0006-9647-2871
Christian Pylatiuk https://orcid.org/0000-0002-3507-7134
Jon Sweeney https://orcid.org/0000-0003-3391-2375

Data availability

All image data used and the code developed for this study are available on Zenodo repository at https://doi.org/10.5281/zenodo.14998760.

References

Ärje J, Melvad C, Jeppesen MR, Madsen SA, Raitoharju J, Rasmussen MS, Iosifidis A, Tirronen V, Gabbouj M, Meissner K, Høye TT (2020) Automatic image-based identification and biomass estimation of invertebrates. Methods in Ecology and Evolution 11: 922–931. https://doi.org/10.1111/2041-210X.13428

Baranchikov Y, Mozolevskaya E, Yurchenko G, Kenis M (2008) Occurrence of the emerald ash borer, *Agrilus planipennis* in Russia and its potential impact on European forestry. EPPO Bulletin 38: 233–238. https://doi.org/10.1111/j.1365-2338.2008.01210.x

Baranchikov Y, Seraya L, Grinash M (2014) All European ash species are susceptible to emerald ash borer *Agrilus planipennis* Fairmaire (Coleoptera: Buprestidae) – a Far Eastern invader. Siberian Journal of Forest Science 6: 80–85.

Brockerhoff EG, Liebhold AM (2017) Ecology of forest insect invasions. Biological Invasions 19: 3141–3159. https://doi.org/10.1007/s10530-017-1514-1

Brydegaard M, Pedales DR, Feng V, Yamoa AS-D, Kouakou B, Månefjord H, Wührl L, Pylatiuk C, Amorim DS, Meier (2024) Towards global insect biomonitoring with frugal methods. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 379(1904): 37920230103. https://doi.org/10.1098/rstb.2023.0103

Cavaletto G, Faccoli M, Marini L, Spaethe J, Magnani G, Rassati D (2020) Effect of trap color on captures of bark-and wood-boring beetles (Coleoptera; Buprestidae and Scolytinae) and associated predators. Insects 11: 749. https://doi.org/10.3390/insects11110749

Chiavassa JA, Kraft M, Noack P, Walther S, Kirse A, Scherber C (2024) The field automatic insect recognition-device — a non-lethal semi-automatic Malaise trap for insect biodiversity monitoring: Proof of concept. Ecology and Evolution 14: e70642. https://doi.org/10.1002/ece3.70642

Crook DJ, Francese JA, Zylstra KE, Fraser I, Sawyer AJ, Bartels DW, Lance DR, Mastro VC (2009) Laboratory and field response of the emerald ash borer (Coleoptera: Buprestidae), to selected regions of the electromagnetic spectrum. Journal of Economic Entomology 102: 2160–2169. https://doi.org/10.1603/029.102.0620

De Cesaro Júnior T, Rieder R (2020) Automatic identification of insects from digital images: A survey. Computers and Electronics in Agriculture 178: 105784. https://doi.org/10.1016/j.compag.2020.105784

Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L (2009) ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Miami, FL, USA, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Dodds KJ, Sweeney J, Francese JA, Besana L, Rassati D (2024) Factors affecting catches of bark beetles and woodboring beetles in traps. Journal of Pest Science 97: 1767–1793. https://doi.org/10.1007/s10340-024-01774-1

Duan JJ, Johnson TD, O'Dea JK, Petrice TR, Haack RA (2024) The ecology, economics, and management of *Agrilus* beetles. Current Forestry Reports 10: 487–509. https://doi.org/10.1007/s40725-024-00230-8

European Food Safety Authority (EFSA), Schans J, Schrader G, Delbianco A, Graziosi I, Vos S (2020a) Pest survey card on *Agrilus planipennis.* EFSA Supporting Publications 17: EN-1945. https://doi.org/10.2903/sp.efsa.2020.EN-1945

European Food Safety Authority (EFSA), Schrader G, Kinkar M, Vos S (2020b) Pest survey card on *Agrilus anxius*. EFSA Supporting Publications 17: EN-1777. https://doi.org/10.2903/sp.efsa.2020.EN-1777

Evans HF, Williams D, Hoch G, Loomans A, Marzano M (2020) Developing a European Toolbox to manage potential invasion by emerald ash borer (*Agrilus planipennis*) and bronze birch borer (*Agrilus anxius*), important pests of ash and birch. Forestry 93: 187–196. https://doi.org/10.1093/forestry/cpz074

Farrugia S (2007) Les *Agrilus* de France (Coleoptera, Buprestidae): une clef de détermination. Magellanes.

Fenn-Moltu G, Ollier S, Caton B, Liebhold AM, Nahrung H, Pureswaran DS, Turner RM, Yamanaka T, Bertelsmeier C (2023) Alien insect dispersal mediated by the global movement of commodities. Ecological Applications 33: e2721. https://doi.org/10.1002/eap.2721

Francese JA, Crook DJ, Fraser I, Lance DR, Sawyer AJ, Mastro VC (2010) Optimization of trap color for emerald ash borer (Coleoptera: Buprestidae). Journal of Economic Entomology 103: 1235–1241. https://doi.org/10.1603/EC10088

Gao Y, Xue X, Qin G, Li K, Liu J, Zhang Y, Li X (2024) Application of machine learning in automatic image identification of insects-a review. Ecological Informatics 80: 102539. https://doi.org/10.1016/j.ecoinf.2024.102539

Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press.

Hansen OL, Svenning JC, Olsen K, Dupont S, Garner BH, Iosifidis A, Price BW, Høye TT (2020) Species-level image classification with convolutional neural network enables insect identification from habitus images. Ecology and Evolution 10: 737–747. https://doi.org/10.1002/ece3.5921

Hartbauer M (2024) Artificial neuronal networks are revolutionizing entomological research. Journal of Applied Entomology 148: 232–251. https://doi.org/10.1111/jen.13227

Imrei Z, Lohonyai Z, Csóka G, Muskovits J, Szanyi S, Vétek G, Fail J, Tóth M, Domingue MJ (2020) Improving trapping methods for buprestid beetles to enhance monitoring of native and invasive species. Forestry 93: 254–264. https://doi.org/10.1093/forestry/cpz071

Isitt R, Liebhold AM, Turner RM, Battisti A, Bertelsmeier C, Blake R, Brockerhoff EG, Heard SB, Krokene P, Økland B, Nahrung HF, Rassati D, Roques A, Yamanaka T, Pureswaran DS (2024) Asymmetrical insect invasions between three world regions. NeoBiota 90: 35–51. https://doi.org/10.3897/neobiota.90.110942

Jendek E, Grebennikov VV (2023) Summary of native geographic distribution of all 3,341 species of the most speciose animal genus *Agrilus* (Coleoptera: Buprestidae). Journal of Insect Biodiversity 39: 32–78. https://doi.org/10.12976/jib/2023.39.2.1

Jocher G, Chaurasia A, Qiu J (2023) Ultralytics YOLO (Version 8.0.0) [Python]. https://github.com/ultralytics/ultralytics

Kelnarova I, Jendek E, Grebennikov VV, Bocak L (2019) First molecular phylogeny of *Agrilus* (Coleoptera: Buprestidae), the largest genus on Earth, with DNA barcode database for forestry pest diagnostics. Bulletin of Entomological Research 109: 200–211. https://doi.org/10.1017/S0007485318000330

Khanam R, Hussain M (2024) YOLOv11: an overview of the key architectural enhancements. arXiv 2410.17725. https://doi.org/10.48550/arXiv.2410.17725

Klooster WS, Gandhi KJ, Long LC, Perry KI, Rice KB, Herms DA (2018) Ecological impacts of emerald ash borer in forests at the epicenter of the invasion in North America. Forests 9: 250. https://doi.org/10.3390/f9050250

Kovacs KF, Haight RG, McCullough DG, Mercader RJ, Siegert NW, Liebhold AM (2010) Cost of potential emerald ash borer damage in US communities, 2009–2019. Ecological Economics 69: 569–578. https://doi.org/10.1016/j.ecolecon.2009.09.004

Kuhn A, San Martin G, Hasbroucq S, Beliën T, Bonte J, Bouget C, Hautier L, Sweeney J, Grégoire JC (2024) Enhancing Buprestidae monitoring in Europe: Trap catches increase with a fluorescent yellow colour but not with the presence of decoys. PLoS ONE 19: e0307397. https://doi.org/10.1371/journal.pone.0307397

Le Souchu E, Bouget C, Sallé A (2024) Environmental drivers of local and temporal variations in the community of oak-associated borers (Coleoptera: Buprestidae). European Journal of Forest Research 143: 603–616. https://doi.org/10.1007/s10342-023-01644-y

Lertrusdachakul I, Leni PE, Gschwind R, Bertheau C (2025) Deep learning for genera-level bark beetle taxonomic classification. Forest Ecology and Management 598: 123190. https://doi.org/10.1016/j.foreco.2025.123190

Limberg C, Melnik A, Harter A, Ritter H (2022) YOLO - you only look 10647 times. Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2023) - Vol. 5: VISAPP, 153–160. https://doi.org/10.48550/arXiv.2201.06159

Loshchilov I, Hutter F (2017) Decoupled weight decay regularization. arXiv:1711.05101. https://doi.org/10.48550/arXiv.1711.05101

Lyal CH, Miller SE (2020) Capacity of United States federal government and its partners to rapidly and accurately report the identity (taxonomy) of non-native organisms intercepted in early detection programs. Biological Invasions 22: 101–127. https://doi.org/10.1007/s10530-019-02147-x

Marais GC, Stratton IC, Johnson AJ, Hulcr J (2025) Progress in developing a bark beetle identification tool. PLoS ONE 20: e0310716. https://doi.org/10.1371/journal.pone.0310716

Muhammad MB, Yeasin M (2020) Eigen-CAM: Class Activation Map using principal components. International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 2020, 1–7. https://doi.org/10.48550/arXiv.2008.00299

Orlova-Bienkowskaja MJ, Drogvalenko AN, Zabaluev IA, Sazhnev AS, Peregudova EY, Mazurov SG, Komarov EV, Struchaev VV, Martynov VV, Nikulina TV, Bieńkowski AO (2020) Current range of *Agrilus planipennis* Fairmaire, an alien pest of ash trees, in European Russia and Ukraine. Annals of Forest Science 77: 29. https://doi.org/10.1007/s13595-020-0930-z

Paiero SM, Jackson MD, Jewiss–Gaines A, Kimoto T, Gill BD, Marshall SA (2012) Field guide to the jewel beetles of Northeastern North America (Coleoptera: Buprestidae). Canadian Food Inspection Agency, 411 pp. https://publications.gc.ca/site/eng/9.699803/publication.html

Poland TM, Petrice TR, Ciaramitaro TM (2019) Trap designs, colors, and lures for emerald ash borer detection. Frontiers in Forests and Global Change 2: 80. https://doi.org/10.3389/ffgc.2019.00080

Pureswaran D, Meurisse N, Rassati D, Liebhold AM, Faccoli M (2022) Climate change and invasion by non-native bark and ambrosia beetles. In: Hofstetter RW, Gandhi K (Eds) Bark beetle management, ecology and climate change. Academic Press, New York, 3–30. https://doi.org/10.1016/B978-0-12-822145-7.00002-7

Rassati D, Marini L, Marchioro M, Rapuzzi P, Magnani G, Poloni R, Di Giovanni F, Mayo P, Sweeney J (2019) Developing trapping protocols for wood-boring beetles associated with broadleaf trees. Journal of Pest Science 92: 267–279. https://doi.org/10.1007/s10340-018-0984-y

Redmon J, Divvala S, Girshick R, Farhadi A (2016) You Only Look Once: unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition, 779–788. https://doi.org/10.1109/CVPR.2016.91

Ruzzier E, Haack RA, Curletti G, Roques A, Volkovitsh MG, Battisti A (2023) Jewels on the go: exotic buprestid around the world (Coleoptera: Buprestidae). NeoBiota 84: 107–135. https://doi.org/10.3897/neobiota.84.90829

Santoiemma G, Battisti A, Courtin C, Curletti G, Faccoli M, Feddern N, Francese J, Franzen EKL, Giannone F, Gossner MM, Kostanowicz C, Marchioro M, Nardi D, Ray AM, Roques A, Sweeney J, Van Rooyen K, Webster V, Rassati D (2024a) Testing a trapping protocol for generic surveillance of wood-boring beetles in heterogeneous landscapes. NeoBiota 95: 77–95. https://doi.org/10.3897/neobiota.95.129483

Santoiemma G, Williams D, Booth EG, Cavaletto G, Connell J, Curletti G, de Groot M, Devine SM, Enston A, Francese JA, Franzen EKL, Giasson M, Groznik E, Gutowski JM, Hauptman T, Hinterstoisser W, Hoch G, Hoppe B, Hughes C, Kostaniwicz C, Peterson D, Plewa R, Ray AM, Sallé A, Sućko K, Sweeney J, Van Rooyen K, Rassati D (2024b) Efficacy of trapping protocols for *Agrilus* jewel beetles (Coleoptera: Buprestidae): a multi-country assessment. Journal of Pest Science 97: 1795–1810. https://doi.org/10.1007/s10340-023-01728-z

Santoiemma G, Sweeney J, Booth EG, Cavaletto G, Curletti G, Devine SM, Francese JA, Franzen EKL, Giannone F, Giasson M, Gutowski JM, Hughes C, Kimoto T, Kostanowicz C, Mokrzycki T, Plewa R, Ray AM, Qingfan M, Williams D, Yan L, Rassati D (2025) Efficacy of unbaited and baited green multi-funnel traps for detection of *Agrilus* species and other wood-boring beetle taxa. Journal of Pest Science 98: 1317–1333. https://doi.org/10.1007/s10340-024-01865-z

Schaefer L (1950) Les Buprestides de France. Miscellanea Entomologica, Supplément, Paris, 511 pp.

Shirali H, Hübner J, Both R, Raupach M, Reischl M, Schmidt S, Pylatiuk C (2024) Image-based recognition of parasitoid wasps using advanced neural networks. Invertebrate Systematics 38: IS24011. https://doi.org/10.1071/IS24011

Silk PJ, Ryall K, Roscoe L (2020) Emerald ash borer, *Agrilus planipennis* (Coleoptera: Buprestidae), detection and monitoring in Canada. Forestry 93: 273–279. https://doi.org/10.1093/forestry/cpz036

Tan M, Le QV (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 6105–6114. https://doi.org/10.48550/arXiv.1905.11946

Tan M, Le QV (2021) EfficientNetV2: Smaller Models and Faster Training. Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 10096–10106. https://doi.org/10.48550/arXiv.2104.00298

Tannous M, Stefanini C, Romano D (2023) A deep-learning-based detection approach for the identification of insect species of economic importance. Insects 14: 148. https://doi.org/10.3390/insects14020148

Teixeira AC, Ribeiro J, Morais R, Sousa JJ, Cunha A (2023) A systematic review on automatic insect detection using deep learning. Agriculture 13: 713. https://doi.org/10.3390/agriculture13030713

Valan M, Makonyi K, Maki A, Vondráček D, Ronquist F (2019) Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks. Systematic Biology 68: 876–895. https://doi.org/10.1093/sysbio/syz014

Volkovitsh MG, Orlova-Bienkowskaja MJ, Kovalev AV, Bieńkowski AO (2020) An illustrated guide to distinguish emerald ash borer (*Agrilus planipennis*) from its congeners in Europe. Forestry 93: 316–325. https://doi.org/10.1093/forestry/cpz024

Wührl L, Pylatiuk C, Giersch M, Lapp F, von Rintelen T, Balke M, Schimdt S, Cerretti P, Meier R (2022) DiversityScanner: Robotic handling of small invertebrates with machine learning methods. Molecular Ecology Resources 22: 1626–1638. https://doi.org/10.1111/1755-0998.13567

Wührl L, Rettenberger L, Meier R, Hartop E, Graf J, Pylatiuk C (2024) Entomoscope: An open-source photomicroscope for biodiversity discovery. IEEE Access 12: 11785–11794. https://doi.org/10.1109/ACCESS.2024.3355272

Supplementary material 1

Supplementary information

Authors: Valerio Caruso, Hossein Shirali, Christophe Bouget, Pierfilippo Cerretti, Gianfranco Curletti, Maarten de Groot, Eva Groznik, Jerzy M. Gutowski, Christian Pylatiuk, Radosław Plewa, Alain Roques, Aurelien Salle, Jon Sweeney, Kate Van Rooyen, Lorenz Wührl, Davide Rassati

Data type: docx

Link: https://doi.org/10.3897/neobiota.105.180959.suppl1